Nucleic Acids Research, 2019 **1**
doi: 10.1093/nar/gkz008

Kinetic analysis of *N*-alkylaryl carboxamide hexitol nucleotides as substrates for evolved polymerases

Marleen Renders[1,†], Shrinivas Dumbre[1,†], Mikhail Abramov[1], Donaat Kestemont[1], Lia Margamuljana[1], Eric Largy[2], Christopher Cozens[3,4], Julie Vandenameele[5], Vitor B. Pinheiro[3,4], Dominique Toye[6], Jean-Marie Frère[5] and Piet Herdewijn[1,7,*]

[1]KU Leuven, Rega Institute for Medical Research, Medicinal Chemistry, Rega, Herestraat 49 box 1041, 3000 Leuven, Belgium, [2]ARNA laboratory, Université de Bordeaux, INSERM U1212, CNRS UMR5320, IECB, 2 rue Robert Escarpit, 33600 Pessac, France, [3]Structural and Molecular Biology Department, University College London, Gower Street, London WC1E B6T, UK, [4]Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, United Kingdom, [5]Laboratory of Enzymology and Protein Folding/Robotein Platform, Centre for Protein Engineering (CIP), Department of Life Sciences, University of Liège, Quartier Agora, Allée du six Août 13, Bât. B6a, 4000 Liège, Belgium, [6]Chemical engineering laboratory, University of Liège, Allée de la chimie, 3, Bât B6c, 4000 Liège, Belgium and [7]Université d'Evry, CNRS-UMR8030/Laboratoire iSSB, CEA, DRF, IG, Genoscope, Université Paris-Saclay, Evry 91000, France

Received November 21, 2018; Revised December 21, 2018; Editorial Decision December 28, 2018; Accepted January 21, 2019

ABSTRACT

Six 1′,5′-anhydrohexitol uridine triphosphates were synthesized with aromatic substitutions appended via a carboxamide linker to the 5-position of their bases. An improved method for obtaining such 5-substituted hexitol nucleosides and nucleotides is described. The incorporation profile of the nucleotide analogues into a DNA duplex overhang using recently evolved XNA polymerases is compared. Long, mixed HNA sequences featuring the base modifications are generated. The apparent binding affinity of four of the nucleotides to the enzyme, the rate of the chemical step and of product release, plus the specificity constant for the incorporation of these modified nucleotides into a DNA duplex overhang using the HNA polymerase T6G12_I521L are determined via pre-steady-state kinetics. HNA polymers displaying aromatic functional groups could have significant impact on the isolation of stable and high-affinity binders and catalysts, or on the design of nanomaterials.

INTRODUCTION

1′,5′-Anhydrohexitol nucleic acids (HNA) have been developed in our laboratory, first for their potential applications in oligonucleotide therapy (1,2) and later for the development of an orthogonal episome for the generation of genetically contained organisms (3–5), for applications in nanotechnology (6,7) and for use in aptamer and aptazyme selections (8,9). The six-membered 1′,5′-anhydrohexitol sugar ring does not contain the glycosidic linkage present in the natural nucleic acids (RNA and DNA), rendering it chemically and enzymatically stable and a noteworthy alternative for the generation of nucleic acid binders, catalysts and nanomaterials (6,8,9). The evolution of DNA-dependent HNA polymerases and reverse transcriptases by Pinheiro *et al.* allowed the sequence-specific synthesis and reverse transcription of HNA fragments necessary for *in vitro* selection experiments and enabled the isolation of functional HNA molecules, including an HNA aptamer against hen egg lysozyme and an endonuclease HNAzyme (8,9). Nonetheless, the negatively charged backbone, together with its general lack of functional groups, limit the binding interactions that HNA (and also the natural nucleic acids, DNA and RNA) can entertain. Base-modified nucleotides have been introduced in selections in order to increase the possible interactions of natural nucleic acids with their binding target (10) or to increase their catalytic potential (11). Due to the synthetic challenge and because of the required high-fidelity and efficient incorporation of the modified nucleotides into a DNA duplex overhang, however, sugar analogues that have modified bases appended

*To whom correspondence should be addressed. Tel: +3216322657; Fax: +3216330026; Email: piet.herdewijn@kuleuven.be
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Present addresses:
Vitor B. Pinheiro, KU Leuven, Rega Institute for Medical Research, Medicinal Chemistry, Rega, Herestraat 49 box 1041, 3000 Leuven, Belgium.
Christopher Cozens, LabGenius Ltd, B201.3 Biscuit Factory, 100 Drummond Road, London SE16 4DG, UK.

onto them have rarely been used for *in vitro* selection strategies (12).

We have synthesized six 1′,5′-anhydrohexitol nucleoside triphosphates with 5-substituted uracil bases involving aromatic residues. These substitutions are linked via a carboxamide group to the uridine nucleotides. The 5-substituted uridine nucleotides can provide an ambiguous hydrogen bonding pattern via the rotation of the exocyclic carbamoyl group, tautomerization, salt concentration (13) and the static interaction of the aryl substituent on the carbamoyl nitrogen can yield a better orientation of the 5-substituent on the base (Figure 1) for increased interactions with a protein target in aptamer selections (14). Carboxamide linkers have been described before for the functionalization at position-5 of the nucleobase and the subsequent isolation of modified nucleic acid aptamers (14). However, in the literature, the synthesis of carbamoyl-modified nucleosides is described starting from 5-iodouracil nucleosides and the conversion to the respective modified compounds is multi-step (15,16) or poor-yielding after Pd (II)-catalyzed carboxyamidation (14,17). Below, we describe an improved chemical pathway to obtain these compounds.

The base substituents that we have synthesized are inspired by the successes that have been obtained using aromatic groups to select high affinity binding aptamers and aptazymes (14,18–27). This is in line with the observation that antibody binding sites often contain a multitude of aromatic residues in their hypervariable domains, critical for recognition (28,29).

Here, via the generation of base-modified HNA nucleotides and their polymerization into highly functionalized HNA sequences, we pave the way for the generation of tight binding, chemically and metabolically stable aptamers. Because the high-fidelity recognition of the modified nucleotides by polymerases is a requirement for their use in *in vitro* selections, the incorporation kinetics of the functionalized hexitol nucleotides by an engineered DNA-dependent HNA polymerase are determined. The synthesis and evaluation of new, hypermodified nucleotides is necessary to advance the field of aptamer therapeutics and diagnostics as only by increasing the number of available aptamer chemistries, we can gain knowledge on the attributes that are essential for success.

MATERIALS AND METHODS

General methods

All oligonucleotides were obtained from Integrated DNA Technologies (Leuven, Belgium) and PAGE purified on a 15 or 20% (depending on the length of the oligo) denaturing gel. Afterwards the oligonucleotides were lyophilized and resuspended in Milli-Q water. ThermoPol® buffer 10× and $MgSO_4$ (100 mM) were purchased from NEB. Ultrapure dTTP, PCR grade, was purchased from Qiagen. Accugel (19:1 acrylamide/bisacrylamide) 40% from National Diagnostics was used to prepare all denaturing polyacrylamide gels. A Typhoon FLA 9500 scanner (GE Healthcare, elongation experiments) was used for the visualization of the PAGE gels containing the fluorescently labelled reactions. The ImageQuant TL™ version 8.1 image analysis software (GE Healthcare) was used for gel band analysis. Graphpad Prism v6 was used for the analysis of the pre-steady-state kinetics, together with Matlab for the global fitting of the generated data.

Elongation reactions

All elongation reactions were carried out on a C1000 Touch Thermal Cycler. The reactions were quenched by addition of gel loading buffer containing 90% formamide, 50 mM EDTA and 0.05% of orange G and heated for 10 min at 95°C before loading onto the gel for analysis.

Annealing of primer and template

The primer was annealed 1:2 to template in 2× ThermoPol® reaction buffer by heating to 95 °C for 5 min and slowly cooling down to room temperature.

Binding affinity determination

The binding affinity between the primer-template complex P3T4 and the enzyme was determined using an octet HTX from Pall – FortéBio. Samples and reagents were placed in black, flat, polypropylene microplates (Greiner 655209), filled with 200 μl/well. Eight channels were used simultaneously, with streptavidin biosensors (Pall – FortéBio 18–5019). Data were recorded using Octet Data Acquisition 8.2.0.9. After biosensors hydration in buffer (89% KB 10×, 10% ThermoPol® buffer 10×, 1% $MgSO_4$ 100 mM), loading of biotinylated primer-template duplex P3T4 at 20 μg/ml was realized during 600 s and remaining accessible sites were quenched during 300 s using biocytin at 10 μg/ml. A baseline in buffer was then recorded during 60 s, followed by 600 s of enzyme association and 600 s of enzyme dissociation in buffer. Five different concentrations of enzyme were used, ranging from 1.56 to 25 nM. Analysis was done with Octet Data Analysis 9.2. The buffer signal was substracted from the sample signals and kinetics were aligned according to the baselines. A global fitting with a 1:1 model was then performed to get $K_{d(DNA)}$, k_{on} and k_{off} values.

Pre-steady-state burst kinetics

All pre-steady-state kinetic experiments were carried out on an SFM-3000 Quench flow instrument from Bio-Logic with the internal temperature maintained at 50°C using a water bath (5 l bath volume) circulating at 10 l/min. The first syringe was filled with enzyme and primer-template in the appropriate concentration in 2× Thermopol buffer. In the second syringe the nucleotide or nucleotide analogue was pipetted, together with $MgSO_4$, yielding an additional 1 mM of Mg^{2+} in the final reaction mixture. The ageing time of the samples varied between 5 ms and 60 s. The samples were quenched with 0.3 M EDTA after which they were mixed with loading dye and analysed on an 18% denaturing PAGE gel. The concentrations of the various modified compounds were as follows: hTTP: 6.25, 12.5, 25, 50 and 75 μM; **1c**, **1d** and **1f**: 6.25, 12.5, 25 and 100 μM; **1e**: 1.6, 6.25, 12.5 and 25 μM. The sampling times were 3, 5, 10, 20, 30 and 60 s for all compounds with the exception of 25 μM **1e** in which case the reaction was also stopped after 0.5, 1.0

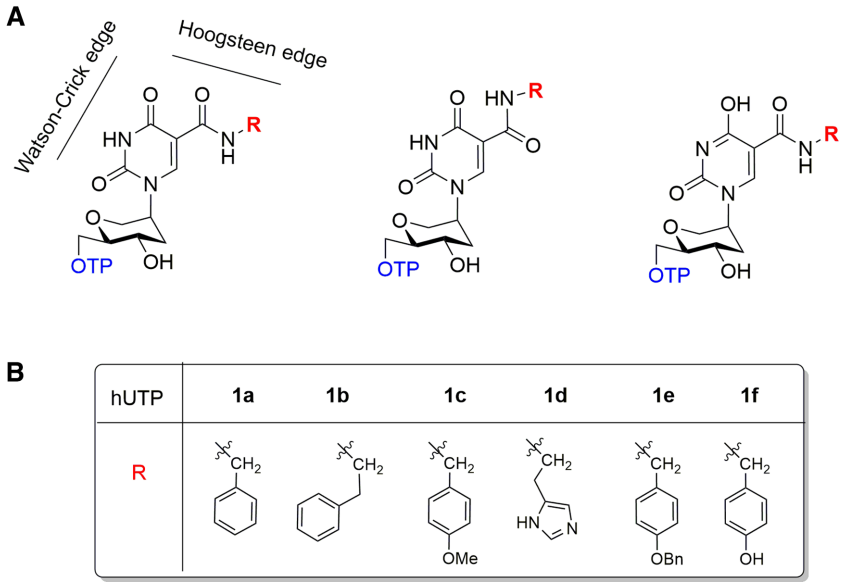


Figure 1. (**A**) Schematic presentation of the modified nucleotides used in this study, **1a–f**, (**B**) the aryl (R) moieties in **1a–f.**

and 2.5 s. When times were larger than 2 s, samples were withdrawn manually. All experiments were carried out in duplicate. The complete set of data were fitted to the Equations (1–4) and (6) using the lsqcurvefit function in Matlab 9.1 and the default optimization parameters, with the two independent variables t and [N] fixed. The initial values for the parameters $K_{d(N)}$, k_2 and k_3 were set to 10 μM, 0.8 s^{-1} and 0.025 s^{-1}. The values for the unknowns were solved by minimizing the difference between the calculated and the experimental values. The final fitted values were checked to not depend on these initial values. The standard deviations were computed from the corresponding diagonal elements of the parameter covariance matrix. For the single burst kinetic experiment as shown in Figure 4, Graphpad Prism v6 was used for data fitting. Non-linear regression curve analysis was applied using least squares fit and initial values of 6.0 for A_0, 0.001 for k_1 and k_2. Again, the final fitted values were checked to not depend on these initial values.

Enzyme active site titration

The enzyme active site was titrated by incubating the enzyme (20 nM concentration as determined by A_{280}) at 50°C with a series of DNA substrate concentrations (between 5 and 100 nM). The reaction was initiated by adding the dTTP substrate at 200 μM and quenched after 50 ms using a quench-flow instrument. The concentrations of product obtained were then fitted to the quadratic equation $[(E_0 + K_{d(DNA)} + [DNA]) - [(E_0 + K_{d(DNA)} + [DNA])^2 - (4E_0[DNA])]^{0.5}]/2$ to give the effective enzyme concentration E_0 and the dissociation constant of the DNA substrate from the polymerase $K_{d(DNA)}$ (30,31).

RESULTS AND DISCUSSION

Synthesis of 5-substituted hexitol uridine nucleotides

Six 5-substituted hUTPs were synthesized as depicted in Scheme 1. The detailed procedures for the synthesis and the analytical data for the compounds and their synthetic intermediates can be found in the Supporting Information. The reaction scheme starts from the commercially available 1,5:2,3-dianhydro-4,6-*O*-benzylidene-D-allitol (2) and 1′,5′-anhydro-4,6-*O*-benzylidene-3-deoxy-D-glucitol (3) building blocks. The regioselective opening of the epoxide moiety in (2) by the DBU salt of 5-iodouracil provides the nucleoside (5) (32). Alternatively, the 3-deoxy analogue of **5** (4) was obtained via the Mitsunobu-type condensation reaction between the alcohol (3) and N^3-benzoyl-5-iodouracil (33). For the generation of the base modifications, molybdenum hexacarbonyl [$Mo(CO)_6$] catalysed (34,35) aminocarbolylation was used. This catalyst itself serves as a source of molecular carbon monoxide allowing an easy reaction execution with 65–92% yield for the 5-substituted uracil nucleosides of 1′,5′-anhydrohexitol (**6a-e**). The xanthate ester of compounds **6a** and **6b** affords the corresponding 3-deoxy-4,6-diols upon Barton–McCombie deoxygenation and acetic acid-mediated benzylidene cleavage. These compounds were treated with $POCl_3$ in trimethylphosphate (TMP) in the presence of Proton-Sponge. Addition of tetrabutylammonium pyrophosphate (TBAPP) directly to the intermediate phosphorodichloridate affords the 5-substituted hUTPs **1a**–**e** in 30–40% yield. The nucleotide **1f** was obtained from nucleotide **1e** by transfer hydrogenolysis catalyzed by palladium on carbon in cyclohexene.

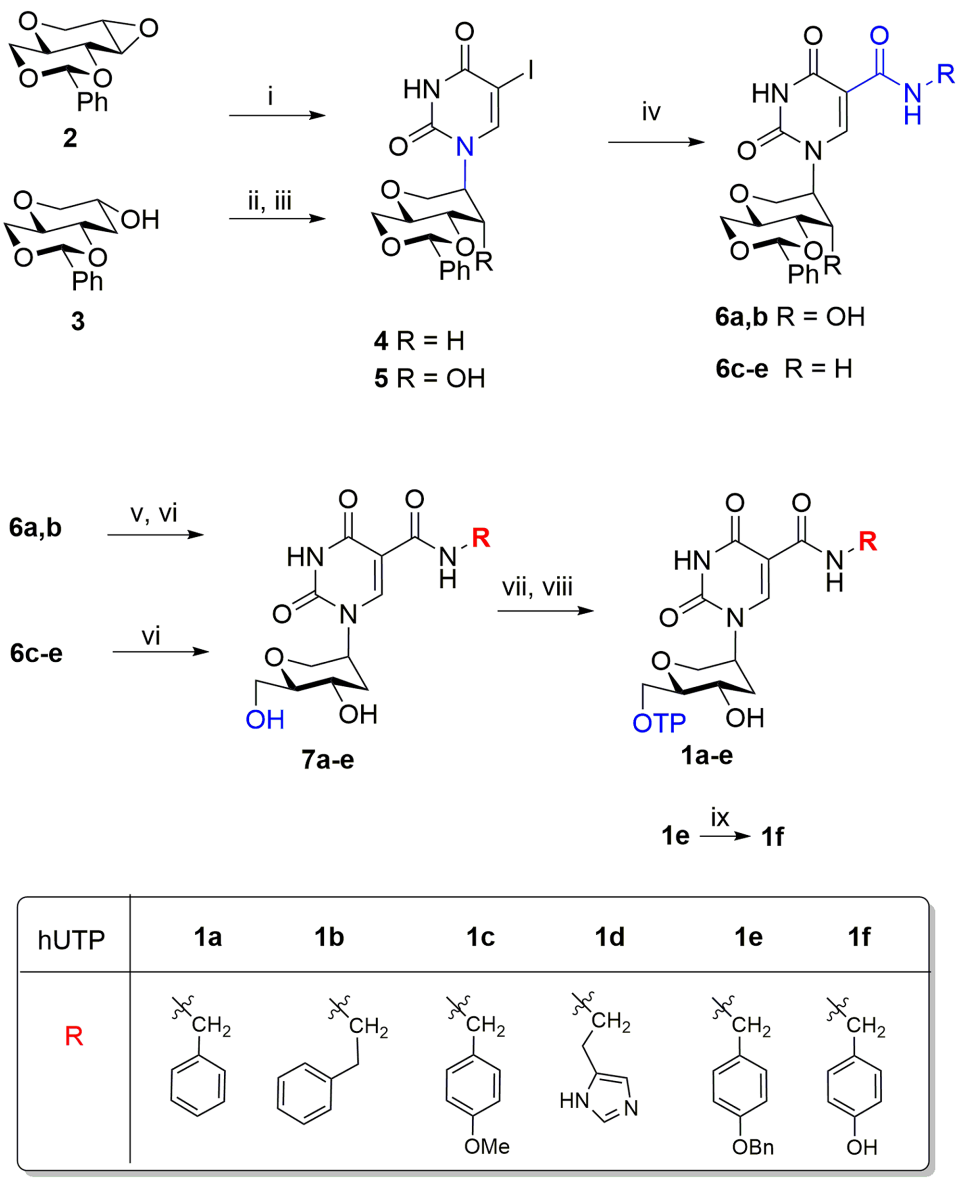
The incorporation of 5-substituted hexitol uracil nucleotides opposite a homopolymeric overhang in a DNA duplex using the engineered HNA polymerases T6G12 and T6G12_I521L

To be able to use a modified nucleotide in an *in vitro* selection, the faithful incorporation of this modified nucleotide into a growing DNA duplex is required. For this reason, we tested whether the six 5-substituted hUTPs are a substrate for polymerases. We used a ten-mer polyA overhang in a DNA primer-template duplex and two variants of the HNA polymerase evolved by Pinheiro *et al.* (8); T6G12, a

Table 1. The oligonucleotide sequences used

Name	Sequence
P1	5′-FAM-CGGATCCGTTTAAGCTAGG-3′
P2	5′-Cy5-CAGGAAACAGCTATGAC-3′
P3	5′-CAGGAAACAGCTATGAC-3′
T1	5′-**GGCCGCAAAAAAAAAA**CCTAGCTTAAACGGATCCG-InvdT-3′
T2	5′-**TGGTCCAGCATCGTGAGATCGATTACCGAACAGCACTACGTGGCTAAGTGCTTATCT**CCTAGCTTAAACGGATCCG-3′
T3	5′-**mUmUmUmUA**GTCATAGCTGTTTCCTG-3′
T4	5′-**CCCCA**GTCATAGCTGTTTCCTG-Biotin-3′

The positions in the sequences where a thymine or uracil nucleotide is expected to be incorporated, are underlined. The template overhangs are highlighted in bold. The template strand T1 is protected at its 3′-end with an inverted dT to prevent 3′-extension of the template strand (36). 2′-*O*-methyl RNA nucleotides are indicated by 'm'.



Scheme 1. Synthesis of the 5-substituted hUTP nucleotides **1a–f**; (i) 5-iodouracil, DBU, CH_3CN, 80°C, 16 h, 63%; (ii) N^3-benzoyliodouracil, PPh_3, DEAD, dioxane, 18 h, RT, 75%; (iii) NH_3, MeOH, 90 min, RT, 65%; (iv) $Mo(CO)_6$, NEt_4Cl, Bu_3N, RNH_2, diglyme, 100–120°C, 2 h, 65–92%; (v) 1,1′-Thiocarbonyldiimidazole, dichloromethane, reflux, 8 h, then Bu_3SnH, AIBN, toluene, 1–1.5 h; 80–82%; (vi) 80% AcOH, 40°C, 1 h, 30–60%; (vii) $POCl_3$, TMP, Proton-Sponge, 0°C, 5 h; (viii) Bu_3N, $(NBu_4)_3HP_2O_7$, DMF, 0–25°C, 30 min, 30–40%; (ix) Pd, cyclohexene.

Thermococcus gorgonarius DNA polymerase variant harbouring 18 mutations compared to the wild type enzyme, and T6G12_I521L, which contains an additional mutation in the palm domain, because these enzymes are specifically evolved to synthesize HNA. The elongation reactions were carried out as described in detail in the Methods section. The primer was 5′-FAM labelled (P1, Table 1) allowing visualization of the elongation products after denaturing PAGE separation. The results are shown in Figure 2.

For both T6G12 and T6G12_I521L, the compounds **1b**, **1c** and **1f** show the best incorporation profiles. Whereas compound **1c** leads to the formation of primarily three products: main product ±1, compounds **1b** and **1f** lead to a ladder of products. With compound **1c** the primer is almost entirely consumed, while with **1b** and **1f** still a significant fraction of the primer remains unreacted.

A 'no template control' reaction was carried out for each of the reactions, to exclude template-independent elongation of the primer with the modified nucleotides (Supplementary Figure S1). This shows that the template-independent extension of the primer is higher with T6G12 than with T6G12_I521L.

Interestingly, when one compares compounds **1a** and **1b**, the extra carbon in **1b** seems to improve the incorporation substantially. Likewise, a substitution in the para-position of the phenylgroup in **1a** with a hydroxyl- (as in **1f**) or a methoxy-functionality (as in **1c**) greatly enhances its capacity as a substrate for the enzymes. A benzyl-functionality (as in **1e**) seems to reduce this beneficial effect on the incorporation, although it still shows an improved incorporation profile compared to **1a**—the compound with the unsubstituted phenylgroup. Histidine-like functionalities have been appended to nucleobases before and the resulting nucleotides are known to be difficult substrates for polymerases (37–40). This seems to remain the case when using evolved polymerases and a hexitol-backbone (**1d**).

The incorporation of 5-substituted hexitol uracil nucleotides opposite a defined mixed sequence overhang in a DNA duplex using *in vitro* evolved polymerases

Given the long-term goal of aptamer/aptazyme selections, we set out to probe the incorporation of the 5-substituted hUTPs into mixed HNA sequences using different mutant polymerases. We tested a series of recently *in vitro* evolved polymerases for this purpose; T6G12, T6G12_I521L, K6G12 - the KOD DNA polymerase variant that has mutations in the same relative positions as the 6G12 TgoT derivative, and TgoT_EPL and TgoT_EPLH - both variants of the TgoT enzyme that were recently described for having improved tPhoNA polymerization activity (8,41). A template (T2, Table 1) was used that is designed to contain all combinations of dinucleotides in such a way that each hN has to be incorporated before and after each of the four hNs at least once, accounting for a 57-mer overhang in total (P1T2 duplex). 5-(*p*-Methoxy)benzylcarboxamide hUTP (**1c**) was chosen for optimization of the reaction conditions. Only with T6G12_I521L in the absence of $MnCl_2$ a defined band of reaction product compatible with full-length incorporation could be obtained. For this reason, we continued with the optimization of the reaction conditions using this enzyme.

The optimal Mg^{2+} concentration for the elongation was determined. 5-(*p*-benzyloxy)benzylcarboxamide hUTP **1e** was chosen for this experiment, because the polymerase stalls the most when incorporating this compound into a mixed sequence (data not shown). The results are shown in Supplementary Figure S3. The elongation seems optimal in the presence of 2 mM extra $MgSO_4$ in addition to the $MgSO_4$ (2 mM) of the Thermopol buffer 1×, i.e. at a concentration of 4 mM $MgSO_4$ final.

Finally, after the optimization of the reaction temperature to cycling at 1 min at 94°C, followed by 5 min at 50°C and 2 h at 65°C, for 16 h, the compounds were incorporated into the P1T2 duplex using T6G12_I521L (1 μM final concentration of enzyme). For all modifications, except for **1e** (truncated product) and **1f** (overshooting) full-length elongation can be obtained (Figure 3). Note that the sequences containing the bulky 5-substituted hU-building blocks are retarded on the gel compared to the HNA sequences without base-substituents. The control reaction in which the template strand is omitted, is shown in Supplementary Figure S4.

Pre-steady-state kinetics of incorporation using T6G12_I521L and the 5-substituted hexitol uracil nucleotides

For an analysis of the incorporation of the heavily modified nucleotides into a growing DNA duplex (P2T3) by the polymerase T6G12_I521L, pre-steady-state kinetics were carried out. The template T3 with the 2′-*O*-methyl RNA overhang immediately after the position where the modified nucleotide has to be incorporated, was chosen for the kinetic studies, to avoid misincorporation of a second modified nucleotide. Given the heavily engineered polymerase that is being used for the incorporation, the likelihood of misincorporation is much higher than with a natural DNA or RNA polymerase.

First, an enzyme (E) active site titration was performed to determine the effective enzyme concentration in the polymerase preparation 'E' of 14.9 ± 0.4 nM (74.5% active fraction) and a binding affinity of the DNA substrate to the polymerase, $K_{d(DNA)}$ of 12.6 nM ± 0.3 nM, which is within the range expected for *wild type* DNA polymerases and very similar to the value obtained using Bio-Layer Interferometry measurements (12.14 ± 0.1 nM) (data not shown) (31,42).

The nucleotide incorporation by a polymerase obeys a rather simple model (43):

$$ED_n + N \underset{K}{\rightleftharpoons} ED_nN \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ED_{n+1} + PP_i$$

$$ED_nN \xrightarrow{k_3} E + D_{n+1}$$

In a first approximation and in order to compare the incorporation of the various compounds, we considered the

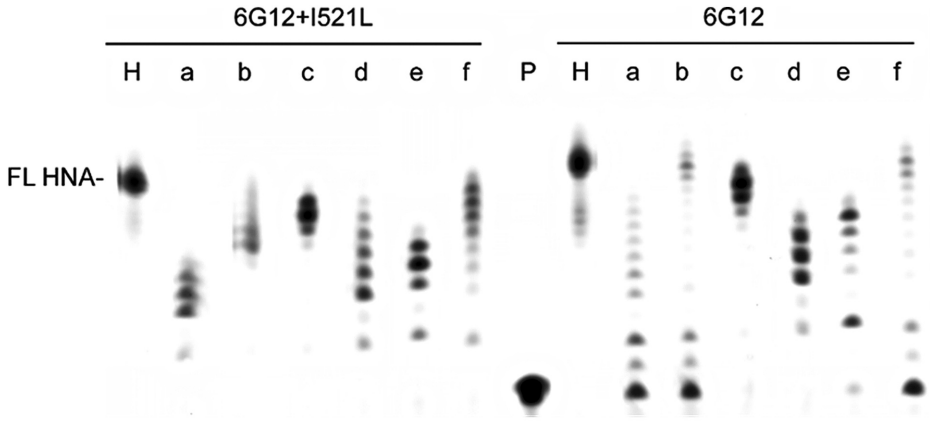


Figure 2. The incorporation of hTTP (H) and the 5-substituted hUTPs **1a-f** (lanes a–f) opposite a poly-dA template overhang in a DNA duplex using the evolved HNA polymerases T6G12_I521L and T6G12. The modified nucleotides (hTTP and **1a**–**f**) are used at a concentration of 125 μM. After optimization of the reaction conditions, the enzymes are used at a final concentration of 51 nM for T6G12_I521L and 82 nM T6G12. The reactions with the enzyme T6G12 contain 0.5 mM freshly prepared $MnCl_2$. All reactions are carried out in 1× Thermopol buffer (NEB) supplemented with 1.5 mM $MgSO_4$. The reactions are incubated at 50°C overnight. The positions where a 5-substituted hUTP has to be incorporated opposite the template oligonucleotide, are underlined in the sequence below the gel image. The lanes indicated with 'P' show the primer control (no enzyme and no nucleotides added). The position for the full-length material of HNA (signifying the incorporation of ten hT nucleotides) is indicated by 'FL HNA' on the side of the gel image.

chemical step (k_2) to be irreversible under our experimental conditions. Indeed, the equilibrium constant of this reaction

$$\frac{[\mathrm{ED}_{n+1}][\mathrm{PPi}]}{[\mathrm{ED}_n N]}$$

can be approximated to 20 M and, with initial nM concentrations of ED_n, the k_{-2} step can be neglected. Our analytical technique measures both the enzyme-bound and the free D_{n+1} so that

$$[\mathrm{P}] = [\mathrm{ED}_{n+1}] + [\mathrm{D}_{n+1}]$$

and, under these conditions,

$$\frac{(P)}{E_0} = A_0\ \left(1 - e^{-\lambda t}\right) + k_{ss}t \qquad (1)$$

In this equation, the first term represents the exponential burst phase (accumulation of ED_{n+1}) and the second one (k_{ss}t) the linear steady-state phase. The values of the three parameters (A_0, λ and k_{ss}) are as follows:

$$A_0 = \frac{k_2'^2}{\left(k_2' + k_3\right)^2} \qquad (2)$$

$$\lambda\ = k_2' + k_3 \qquad (3)$$

$$k_{ss} = \frac{k_2' k_3}{k_2' + k_3} \qquad (4)$$

and

$$\frac{k_{ss}}{A_0} = \frac{k_3\left(k_2' + k_3\right)}{k_2'} \qquad (5)$$

where

$$k_2' = \frac{k_2\,[N]}{K_{d(N)} + [N]} \qquad (6)$$

For instance, Figure 4 shows the time-course of P accumulation when $[E]_0$ is 14.9 nM, $[D_n]_0 = 100$ nM and $[N] = 75$ μM. From the $[P]/[E]_0$ versus time curve, the following values can be deduced using fitting of Equation (1) as described in the Materials and Methods section:

$$\lambda = (0.58 \pm 0.02)\, s^{-1}$$

$$\mathrm{A_0} = 0.96 \pm 0.02$$

$$k_{ss} = (0.025 \pm 0.001)\, \mathrm{s}^{-1}$$

In a first approximation, good estimations would be 0.58 s^{-1} and 0.026 s^{-1} for the values of k'_2 and k_3, respectively.

We verified that the amplitude of the burst phase increased when the enzyme concentration increased and that all modified nucleotides in this study exhibited an initial burst in their incorporation profiles. This indicates that a step after the chemical one (e.g. DNA product dissociation from the enzyme) is rate-limiting for product formation in these modified nucleotide incorporations. Additionally, we controlled the saturation of the enzyme with the DNA substrate at the concentrations used (data not shown) (44,45).

Finally, the burst kinetics were determined for a range of nucleotide concentrations for hTTP and the modified nucleotides **1c**–**1f**. The results are shown in Table 2, where k_{pol} represents the maximum rate of incorporation and $K_{d(N)}$ the dissociation constant of the complex ED_nN. The large errors (particularly for **1c**) are due to the relatively low number of data points (due in turn to the limited amounts of modified nucleotides available) and to the fact that, in most

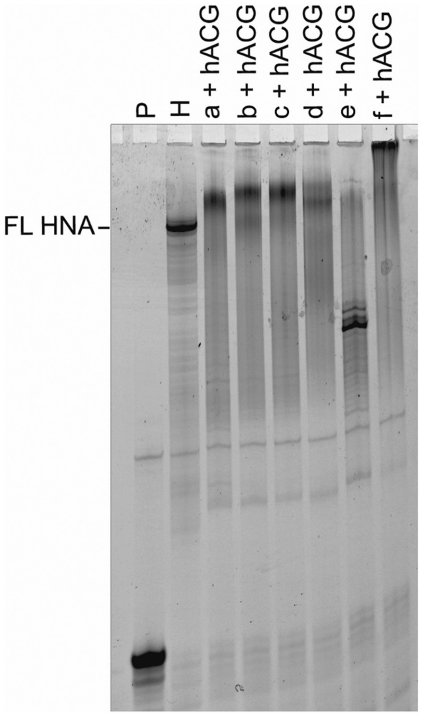


Figure 3. The incorporation of the 5-substituted hUTPs together with hATP, hCTP and hGTP into the P1T2 duplex overhang using T6G12_I 521L after cycling for 1 min at 94°C, followed by 5 min at 50°C and 2 h 65°C, for 16 h in total, in Thermopol buffer 1× containing an additional 2 mM $MgSO_4$. 'P' indicates the primer control. Lane H shows the hNTP control. Lanes a + hACG to f + hACG show the incorporation of the hA, hC and hG building blocks together with **1a**–**f** respectively.

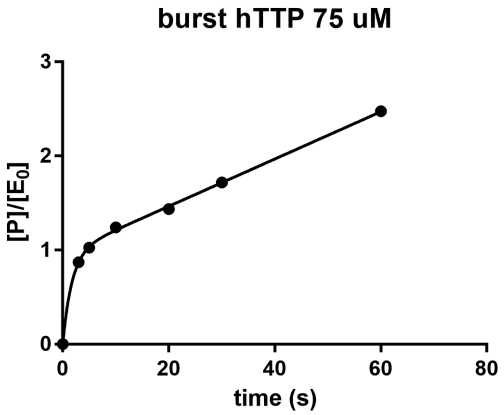


Figure 4. The burst kinetic profile of the incorporation of hTTP at a concentration of 75 μM into the growing duplex P2T3 (100 nM) using T6G12_I521L at a concentration of 14.9 nM as determined via the titration of the enzyme active site. Graphpad Prism v6 was used for data fitting as described in the Materials and Methods section (R^2 value of 0.997).

cases, the samples were manually pipetted, resulting in a relatively large error on the sampling time. Despite the very large errors recorded for **1c**, it is interesting to note that the $k_{pol}/K_{d(N)}$ ratio is similar to that of hTPP. Note that, for compounds **1a** and **1b**, small levels of misincorporation or primer degradation were observed, preventing the determination of the incorporation parameters.

The deduced k_3 values for compounds the **1d** and **1e** are significantly lower than for the three other ones. This is surprising since one would not expect such a large influence of the added nucleotide on the dissociation rate of the ED_{n+1} complex. Moreover, in some cases, the k_{ss} values, when determined directly from the linear portions of the curves could be as high as 0.01 s^{-1} (not shown). Since $k_{ss} \leq k_3$ (Equation 4), the fitting procedure was repeated for **1d** and **1e** with fixed k_3 values of 0.01 and 0.02 s^{-1}. The value of 0.02 s^{-1} for the dissociation of the ED_n complex has been described before in the literature (43). The results are also shown in Table 2 and indicate that the value of k_3 has only a minor influence on those of the other two parameters.

The ratio $k_{pol}/K_{d(N)}$ represents the specificity constant for the incorporation during the processive synthesis, a more accurate estimation than k_{cat}/K_m obtained under steady-state conditions. Globally, our results show that all the modified nucleotides synthesized in this study can be incorporated into a DNA duplex with $k_{pol}/K_{d(N)}$ values similar to that of hTTP. The only exception is the imidazole-substituted **1d** for which the second-order rate constant is at least one order of magnitude lower than those of the other compounds.

Although the $K_{d(N)}$s in Table 2 are within the expected ranges for DNA polymerases, the incorporation rates (k_{pol}) are lower than would be obtained with natural thermophilic polymerases. This is likely the collateral damage caused by the mutations that are needed to push the enzyme towards HNA polymerization. Especially for the compounds **1d** and **1f** the incorporation rates are somewhat lower than that for an hT building block.

CONCLUSION

We have successfully synthesized 5-substituted hexitol uracil nucleotides, using an improved chemical synthesis pathway compared to the previously described methods towards obtaining 5-substituted nucleotides. We have then used the modified nucleotides for polymerizing HNA sequences with aromatic substituents on the U-bases. Such polymers might be important tools in the selections of aptamers to difficult protein or small molecule targets, as has been seen for DNA (18,21,24,46). The kinetic parameters for the incorporation of four of the modified nucleotides using an *in vitro* evolved HNA polymerase were determined. Despite the quite large errors on the experimental values due to manual pipetting of the samples within a short time frame, the data are sufficient to conclude that, with the exception of compound **1d**, all compounds have a similar incorporation rate using the evolved enzyme and the base substitutions do not have a large influence on the specificity of the incorporation. Although the apparent binding affinities of the enzyme for the 5-substituted hUTPs are within the range expected for thermophilic nucleic acid polymerases, the incorporation rates seem markedly low. This indicates that, although the mutated polymerase has been evolved towards HNA polymerization and does accept the 5-substituted hexitol building blocks as a substrate, the incorporation is impeded. This information is important for future *in vitro* evolution experiments. The mutations that have been introduced in the evolved polymerase

Table 2. The values of the kinetic constants for the incorporation of hTTP and the nucleotide analogues **1c**–**1f** into the growing DNA duplex P2T3 using T6G12_I521L deduced by global fitting using Matlab 9.1 (least square method)

	$k_2 = k_{pol}$ (s^{-1})	$K_{d(N)}$ (μM)	$k_{pol}/K_{d(N)}$ (mM^{-1} s^{-1})	k_3 (s^{-1})
hTTP	0.62 ± 0.13	22.0 ± 8	29	0.028 ± 0.002
1c	2 ±1.8	91 ± 90	22	0.026 ± 0.002
1d	0.09 ± 0.05	38 ± 6.3	2.4	0.002 ± 0.0015
1d	0.07 ± 0.04	32 ± 5	2.2	**0.01**
1d	0.05 ± 0.005	24 ± 5	2.1	**0.02**
1e	0.29 ± 0.04	1.2 ± 0.6	240	0.004 ± 0.001
1e	0.25 ± 0.05	2.4 ± 1.2	104	**0.01**
1e	0.15 ± 0.04	3.6 ± 2	42	**0.02**
1f	0.12 ± 0.01	2.4 ± 0.2	50	0.015 ± 0.001

For **1d** and **1e**, the fitting was repeated with fixed values of k_3 (0.01 and 0.02 s^{-1}, as indicated in bold).
The standard deviations were computed from the corresponding diagonal elements of the parameter covariance matrix. An example of parity plot is given in the SI (Supplementary Figure S5).

can mainly be found in the thumb region, which clamps the DNA down in the enzyme, near the active site, and plays an important role in the translocation and the processivity of the enzyme. We previously thought the detrimental factor in the polymerase mechanism, yielding poor polymerization, was the unsuccessful translocation of the enzyme after the incorporation of a modified nucleotide. The kinetic studies performed here show that the polymerase is also lacking in speed when it comes to the chemical step of the reaction. Future evolution experiments towards improving these polymerases might, for this reason, benefit from mutations focussed in the region surrounding (if not immediately in) the active site (47). The evolution of efficient and accurate HNA polymerases is essential to the advancement of this field of research.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Special thanks to Chantal Biernaux for editorial help.

FUNDING

European Research Council under the European Union's Seventh Framework Program (FP7/2007–2013)/ERC [ERC-2012-ADG 20120216/ 320683 to P.H.]; FWO Flanders Research Foundation [1247114N to M.R.]; Walloon Region (SPW, DGO6, Belgium) for supporting the ROBOTEIN project (2014–2016) within the frame of the EQUIP2013 program [convention 1318159]. The salary of Julie Vandenameele is currently supported by the AUTOBMP2 project [SPW, DGO6, convention 16100518]. Funding for open access charge: Institutional budget.
Conflict of interest statement. None declared.

REFERENCES

1. Kang,H., Fisher,M.H., Xu,D., Miyamoto,Y.J., Marchand,A., Van Aerschot,A., Herdewijn,P. and Juliano,R.L. (2004) Inhibition of MDR1 gene expression by chimeric HNA antisense oligonucleotides. *Nucleic Acids Res.*, **32**, 4411–4419.
2. Le,B.T., Chen,S., Abramov,M., Herdewijn,P. and Veedu,R.N. (2016) Evaluation of anhydrohexitol nucleic acid, cyclohexenyl nucleic acid and d-altritol nucleic acid-modified 2′-O-methyl RNA mixmer antisense oligonucleotides for exon skipping in vitro. *Chem. Commun. (Camb.)*, **52**, 13467–13470.
3. Pezo,V., Liu,F.W., Abramov,M., Froeyen,M., Herdewijn,P. and Marliere,P. (2013) Binary genetic cassettes for selecting XNA-templated DNA synthesis in vivo. *Angew. Chem. Int. Ed. Engl.*, **52**, 8139–8143.
4. Pezo,V., Schepers,G., Lambertucci,C., Marliere,P. and Herdewijn,P. (2014) Probing ambiguous base-pairs by genetic transformation with XNA templates. *ChemBioChem.*, **15**, 2255–2258.
5. Pochet,S., Kaminski,P.A., Van Aerschot,A., Herdewijn,P. and Marlière,P. (2003) Replication of hexitol oligonucleotides as a prelude to the propagation of a third type of nucleic acid in vivo. *C. R. Biol.*, **326**, 1175–1184.
6. Taylor,A.I., Beuron,F., Peak-Chew,S.Y., Morris,E.P., Herdewijn,P. and Holliger,P. (2016) Nanostructures from synthetic genetic polymers. *ChemBioChem.*, **17**, 1107–1110.
7. Pinheiro,V.B. and Holliger,P. (2014) Towards XNA nanotechnology: new materials from synthetic genetic polymers. *Trends Biotechnol.*, **32**, 321–328.
8. Pinheiro,V.B., Taylor,A.I., Cozens,C., Abramov,M., Renders,M., Zhang,S., Chaput,J.C., Wengel,J., Peak-Chew,S.Y., McLaughlin,S.H. *et al.* (2012) Synthetic genetic polymers capable of heredity and evolution. *Science*, **336**, 341–344.
9. Taylor,A.I., Pinheiro,V.B., Smola,M.J., Morgunov,A.S., Peak-Chew,S., Cozens,C., Weeks,K.M., Herdewijn,P. and Holliger,P. (2015) Catalysts from synthetic genetic polymers. *Nature*, **518**, 427–430.
10. Diafa,S. and Hollenstein,M. (2015) Generation of aptamers with an expanded chemical repertoire. *Molecules*, **20**, 16643–16671.
11. Hollenstein,M. (2015) DNA Catalysis: The chemical repertoire of DNAzymes. *Molecules*, **20**, 20777–20804.
12. Mei,H., Liao,J.Y., Jimenez,R.M., Wang,Y., Bala,S., McCloskey,C., Switzer,C. and Chaput,J.C. (2018) Synthesis and evolution of a threose nucleic acid aptamer bearing 7-Deaza-7-Substituted guanosine residues. *J. Am. Chem. Soc.*, **140**, 5706–5713.
13. Park,C. and Raines,R.T. (2001) Quantitative analysis of the effect of salt concentration on enzymatic catalysis. *J. Am. Chem. Soc.*, **123**, 11472–11479.
14. Vaught,J.D., Bock,C., Carter,J., Fitzwater,T., Otis,M., Schneider,D., Rolando,J., Waugh,S., Wilcox,S.K. and Eaton,B.E. (2010) Expanding the chemistry of DNA for in vitro selection. *J. Am. Chem. Soc.*, **132**, 4141–4151.
15. El Safadi,Y., Paillart,J.C., Laumond,G., Aubertin,A.M., Burger,A., Marquet,R. and Vivet-Boudou,V. (2010) 5-Modified-2′-dU and 2′-dC as mutagenic anti HIV-1 proliferation agents: synthesis and activity. *J. Med. Chem.*, **53**, 1534–1545.
16. Nomura,Y., Ueno,Y. and Matsuda,A. (1997) Site-specific introduction of functional groups into phosphodiester oligodeoxynucleotides and their thermal stability and nuclease-resistance properties. *Nucleic Acid Res.*, **25**, 2784–2791.
17. Bhanage,B., Tambade,P., Patil,Y. and Bhanushali,M. (2008) Pd(OAc)2-Catalyzed aminocarbonylation of Aryl iodides with aromatic or aliphatic amines in water. *Synthesis*, **2008**, 2347–2352.

18. Gold,L., Ayers,D., Bertino,J., Bock,C., Bock,A., Brody,E.N., Carter,J., Dalby,A.B., Eaton,B.E., Fitzwater,T. *et al.* (2010) Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*, **5**, e15004.
19. Cox,J. and Mann,M. (2011) Quantitative, high-resolution proteomics for data-driven systems biology. *Annu. Rev. Biochem.*, **80**, 273–299.
20. Gelinas,A.D., Davies,D.R., Edwards,T.E., Rohloff,J.C., Carter,J.D., Zhang,C., Gupta,S., Ishikawa,Y., Hirota,M., Nakaishi,Y. *et al.* (2014) Crystal structure of interleukin-6 in complex with a modified nucleic acid ligand. *J. Biol. Chem.*, **289**, 8720–8734.
21. Davies,D.R., Gelinas,A.D., Zhang,C., Rohloff,J.C., Carter,J.D., O'Connell,D., Waugh,S.M., Wolk,S.K., Mayfield,W.S., Burgin,A.B. *et al.* (2012) Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *PNAS*, **109**, 19971–19976.
22. Gupta,S., Hirota,M., Waugh,S.M., Murakami,I., Suzuki,T., Muraguchi,M., Shibamori,M., Ishikawa,Y., Jarvis,T.C., Carter,J.D. *et al.* (2014) Chemically modified DNA aptamers bind interleukin-6 with high affinity and inhibit signaling by blocking its interaction with interleukin-6 receptor. *J. Biol. Chem.*, **289**, 8706–8719.
23. Hopfield,J.J. (1974) Kinetic Proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 4135–4139.
24. Hathout,Y., Brody,E., Clemens,P.R., Cripe,L., DeLisle,R.K., Furlong,P., Gordish-Dressman,H., Hache,L., Henricson,E., Hoffman,E.P. *et al.* (2015) Large-scale serum protein biomarker discovery in Duchenne muscular dystrophy. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 7153–7158.
25. Rohloff,J.C., Gelinas,A.D., Jarvis,T.C., Ochsner,U.A., Schneider,D.J., Gold,L. and Janjic,N. (2014) Nucleic acid ligands with Protein-like side Chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther. Nucleic Acids*, **3**, e201.
26. Tarasow,T.M., Tarasow,S.L. and Eaton,B.E. (1997) RNA-catalysed carbon-carbon bond formation. *Nature*, **389**, 54–57.
27. Wiegand,T.W., Janssen,R.C. and Eaton,B.E. (1997) Selection of RNA amide synthases. *Chem. Biol.*, **4**, 675–683.
28. Chen,Y., Wiesmann,C., Fuh,G., Li,B., Christinger,H.W., McKay,P., de Vos,A.M. and Lowman,H.B. (1999) Selection and analysis of an optimized anti-VEGF antibody: crystal structure of an affinity-matured Fab in complex with antigen. *J. Mol. Biol.*, **293**, 865–881.
29. Fellouse,F.A., Wiesmann,C. and Sidhu,S.S. (2004) Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 12467–12472.
30. Johnson,K.A. (1992) 1 Transient-State kinetic analysis of enzyme reaction pathways. **20**, 1–61.
31. Johnson,K.A. (1995) Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. *Methods Enzymol.*, **249**, 38–61.
32. Abramov,M. and Herdewijn,P. (2007) Synthesis of altritol nucleoside phosphoramidites for oligonucleotide synthesis. *Curr. Protoc. Nucleic Acid Chem.*, doi:10.1002/0471142700.nc0118s30.
33. Ostrowski,T., Wroblowski,B., Busson,R., Rozenski,J., De Clercq,E., Bennett,M.S., Champness,J.N., Summers,W.C., Sanderson,M.R. and Herdewijn,P. (1998) 5-Substituted pyrimidines with a 1,5-anhydro-2, 3-dideoxy-D-arabino-hexitol moiety at N-1: synthesis, antiviral activity, conformational analysis, and interaction with viral thymidine kinase. *J. Med. Chem.*, **41**, 4343–4353.
34. Ren,W. and Yamane,M. (2010) Mo(CO)(6)-mediated carbamoylation of aryl halides. *J. Org. Chem.*, **75**, 8410–8415.
35. Wannberg,J. and Larhed,M. (2003) Increasing rates and scope of reactions: sluggish amines in microwave-heated aminocarbonylation reactions under air. *J. Org. Chem.*, **68**, 5750–5753.
36. Rösch,H., Fröllich,A., Ortigao,J.F.R., Montenarh,M. and Seliger,H. (1990) Patent US US5750669A.
37. Hipolito,C.J., Hollenstein,M., Lam,C.H. and Perrin,D.M. (2011) Protein-inspired modified DNAzymes: dramatic effects of shortening side-chain length of 8-imidazolyl modified deoxyadenosines in selecting RNaseA mimicking DNAzymes. *Org. Biomol. Chem.*, **9**, 2266–2273.
38. Perrin,D.M., Garestier,T. and Hélène,C. (1999) Expanding the catalytic repertoire of nucleic acid catalysts: simultaneous incorporation of two modified deoxyribonucleoside triphosphates bearing ammonium and imidazolyl functionalities. *Nucleosides Nucleotides*, **18**, 377–391.
39. Gourlain,T., Sidorov,A., Mignet,N., Thorpe,S.J., Lee,S.E., Grasby,J.A. and Williams,D.M. (2001) Enhancing the catalytic repertoire of nucleic acids. II. Simultaneous incorporation of amino and imidazolyl functionalities by two modified triphosphates during PCR. *Nucleic Acids Res.*, **29**, 1898–1905.
40. Santoro,S.W., Joyce,G.F., Sakthivel,K., Gramatikova,S. and Barbas,C.F. (2000) RNA cleavage by a DNA enzyme with extended chemical functionality. *J. Am. Chem. Soc.*, **122**, 2433–2439.
41. Liu,C., Cozens,C., Jaziri,F., Rozenski,J., Marechal,A., Dumbre,S., Pezo,V., Marliere,P., Pinheiro,V.B., Groaz,E. *et al.* (2018) Phosphonomethyl oligonucleotides as Backbone-Modified artificial genetic polymers. *J. Am. Chem. Soc.*, **140**, 6690–6699.
42. Tsai,Y.C. and Johnson,K.A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry*, **45**, 9675–9687.
43. Estep,P.A. and Johnson,K.A. (2011) Effect of the Y955C mutation on mitochondrial DNA polymerase nucleotide incorporation efficiency and fidelity. *Biochemistry*, **50**, 6376–6386.
44. Gardner,A.F., Joyce,C.M. and Jack,W.E. (2004) Comparative kinetics of nucleotide analog incorporation by vent DNA polymerase. *J. Biol. Chem.*, **279**, 11834–11842.
45. Schermerhorn,K.M. and Gardner,A.F. (2015) Pre-steady-state kinetic analysis of a family D DNA polymerase from thermococcus sp. 9 degrees N reveals mechanisms for archaeal genomic replication and maintenance. *J. Biol. Chem.*, **290**, 21800–21810.
46. Rohloff,J.C., Gelinas,A.D., Jarvis,T.C., Ochsner,U.A., Schneider,D.J., Gold,L. and Janjic,N. (2014) Nucleic acid ligands with Protein-like side Chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther. Nucleic Acids*, **3**, e201.
47. Cozens,C., Pinheiro,V.B., Vaisman,A., Woodgate,R. and Holliger,P. (2012) A short adaptive path from DNA to RNA polymerases. *PNAS*, **109**, 8067–8072.